The Origin of Consciousness

By Richard Preschel

Abstract: This article examines a radical scientific hypothesis proposing that consciousness originates in mitochondria rather than the brain's neural networks. By analyzing how anesthetics disrupt the electrochemical gradients within these organelles across various life forms, the author suggests that even simple organisms like amoebas possess a primitive form of awareness. This "proton-motive force" serves as a two-billion-year-old biological substrate that represents an organism's internal state through pure electrochemistry. The text argues that feeling is essentially metabolism viewed from an internal perspective, making it a universal characteristic of cellular life. Ultimately, the source concludes that while brains refined the experience of consciousness, they merely inherited its ancient mitochondrial foundation.

There is a question that has haunted philosophy since at least the ancient Greeks, and that modern neuroscience, for all its sophistication, has barely touched: where does consciousness begin? Not how it works in the human brain — that is a hard enough problem — but when, in the history of life, did anything begin to feel anything at all?

The standard answer, rarely examined, is that consciousness arrived with nervous systems. Complex brains produce complex experience; simpler brains produce simpler experience; and below a certain threshold of

neural organisation, there is nothing it is like to be alive. A bacterium, on this view, is a biological machine operating entirely in the dark. An insect is probably dark too. A fish, debatable. A dog, most likely yes.

This answer is comfortable because it confines consciousness to a familiar neighbourhood. But it has a problem: it is not really an answer at all. It simply relocates the mystery. Why should neurons produce experience when other cells do not? What is it about a synapse that generates feeling, when the molecules doing the work — ion channels, membrane gradients, electrochemical potentials — are present, in various forms, in virtually every living cell? The neurocentric view doesn't explain consciousness. It merely draws a line and calls everything above it conscious, without explaining what the line means.

Nick Lane, Professor of Evolutionary Biochemistry at University College London, has proposed something more radical and, I will argue, more honest: that consciousness does not begin with neurons. It begins with mitochondria. And the argument, followed carefully, leads somewhere few biologists have been willing to go.

Consciousness Is Soluble in Ether

The starting point is a provocation borrowed from the biophysicist Luca Turin: almost the only thing we know for certain about consciousness is that it is, so to speak, soluble in ether, chloroform, and a variety of other solvents. When you inhale chloroform, consciousness dissolves. When the drug clears, it returns. Whatever consciousness is, anesthetics can abolish it — reliably, reversibly, and in every animal in which the experiment has been tried.

This fact is more puzzling than it appears. Anesthetics are a chemically diverse group. Ether, chloroform, nitrous oxide, xenon — a noble gas — and various steroids all produce general anesthesia, yet share no obvious structural similarity. For decades, the standard explanation was that they

dissolve into the lipid membranes of neurons, disrupting their function. But this explanation has never been fully satisfying, not least because neurons are not the only cells with lipid membranes.

More recently, Lane and collaborators have pursued a different lead. Anesthetics, it turns out, affect mitochondria. Specifically, they appear to disrupt a quantum-mechanical process called **chiral-induced spin selectivity** (CISS) — the polarisation of electron spin that occurs as electrons flow through the respiratory chain of the mitochondrial inner membrane. The anesthetic doesn't simply knock out neurons. It interferes with something far older and more fundamental: the flow of electrons through the molecular machinery that powers all complex life.

And crucially, this effect is not confined to animals with nervous systems. Anesthetics render amoebas immobile. They suppress the behaviour of organisms with no neurons whatsoever. Which raises a question that Lane takes seriously and that mainstream neuroscience has largely evaded: if you can make an amoeba unconscious, was it conscious before?

The Machinery of Life and the Origin of Feeling

To understand what Lane is proposing, you need to understand what mitochondria actually do — and how extraordinary the process is.

As the biologist Lynn Margulis demonstrated in a paper that was rejected by fifteen journals before it changed biology forever, mitochondria are the descendants of free-living bacteria that were engulfed by a host cell roughly two billion years ago, in an event so improbable that it appears to have happened only once in the history of life on Earth. They retain their own genome, their own membranes, their own ancient biochemistry. And at the core of that biochemistry is one of the most remarkable physical processes in nature.

The inner mitochondrial membrane — folded into elaborate structures

called cristae — is the site of the **electron transport chain**: a series of large protein complexes through which electrons, stripped from food molecules by the Krebs cycle, are passed in sequence. At the end of this chain, at a complex called cytochrome c oxidase, the electrons are handed to oxygen, which combines with protons to form water. This is why we breathe: oxygen is the final acceptor in a chain of electron transfers that began with what we ate.

The energy released as electrons move through this chain is used to do something elegantly simple: pump protons — hydrogen ions — from the interior of the mitochondrion across the inner membrane to the intermembrane space. This creates a gradient: more protons outside than inside, and a positive electrical charge outside relative to inside. The resulting force — the **proton-motive force** — is the primary energy currency of the cell, more fundamental even than ATP.

ATP, the cell's rechargeable energy currency, is made when those protons flood back through the only gate available to them: a molecular turbine called **ATP synthase**, which physically rotates as protons pass through it, using the rotation to bond phosphate groups onto ADP. The result is ATP — the molecule that powers virtually every cellular process, from muscle contraction to protein synthesis to the firing of neurons.

The voltage across the inner mitochondrial membrane is approximately 150 to 200 millivolts. That sounds modest. But the membrane is only five nanometres thick — and an electric field of that magnitude across a layer that thin works out to roughly 30 million volts per metre. Sustained, continuously, across the inner membrane of every mitochondrion in every cell of your body, right now, as you read this.

This is the physical state that Lane proposes as the substrate of the most primitive feeling. The proton-motive force is not merely an energy-storage mechanism. It is a moment-by-moment electrical representation of how the cell is faring in relation to its environment. When nutrients are plentiful and oxygen flows freely, the gradient is steep and stable. When the cell is starved, poisoned, or stressed, the

gradient falters. The cell's own energy state encodes its condition — continuously, physically, in the language of electrochemistry.

Lane's hypothesis is that this fluctuating electrochemical state *is* the most primitive feeling. Not a metaphor for feeling. Not an analogy. The actual physical substrate of what, in its most elaborate form, becomes human emotion, pain, pleasure, and self-awareness. A continuous signal, present in every cell that respires, two billion years older than the first neuron.

Why Feelings Must Be Physical — and Old

The argument is not merely poetic. It is grounded in evolutionary logic that Lane, as an evolutionary biochemist, takes with full seriousness.

If feelings are real — if they genuinely influence what organisms do — then they must be subject to natural selection. Anything that affects behaviour and has a physical basis can be selected for or against. But this means feelings cannot be epiphenomenal ghosts hovering above the biology. They must be physical states, measurable in principle even if not yet in practice, that make a difference to what an organism does.

The most primitive feeling Lane identifies is binary and stark: something like *I am doing well* versus *I am not doing well*. Not thought, not reflection, not the rich texture of human experience — just a raw signal of biological status. And this signal, he argues, must be as old as the capacity for directed response to environment. An organism that can register its own condition and act on that registration has something natural selection can work with. An organism that cannot is entirely blind to itself.

Feelings such as pain, in this view, are not primarily about information processing in neural networks. Pain has very little cognitive content. It is not a thought. It is a state — urgent, insistent, biological. Lane's proposal is that this state has its roots not in the nervous system but in the ancient electrochemistry of cellular respiration: the same proton gradient that powers the amoeba, the bacterium, the plant cell, and the neuron alike.

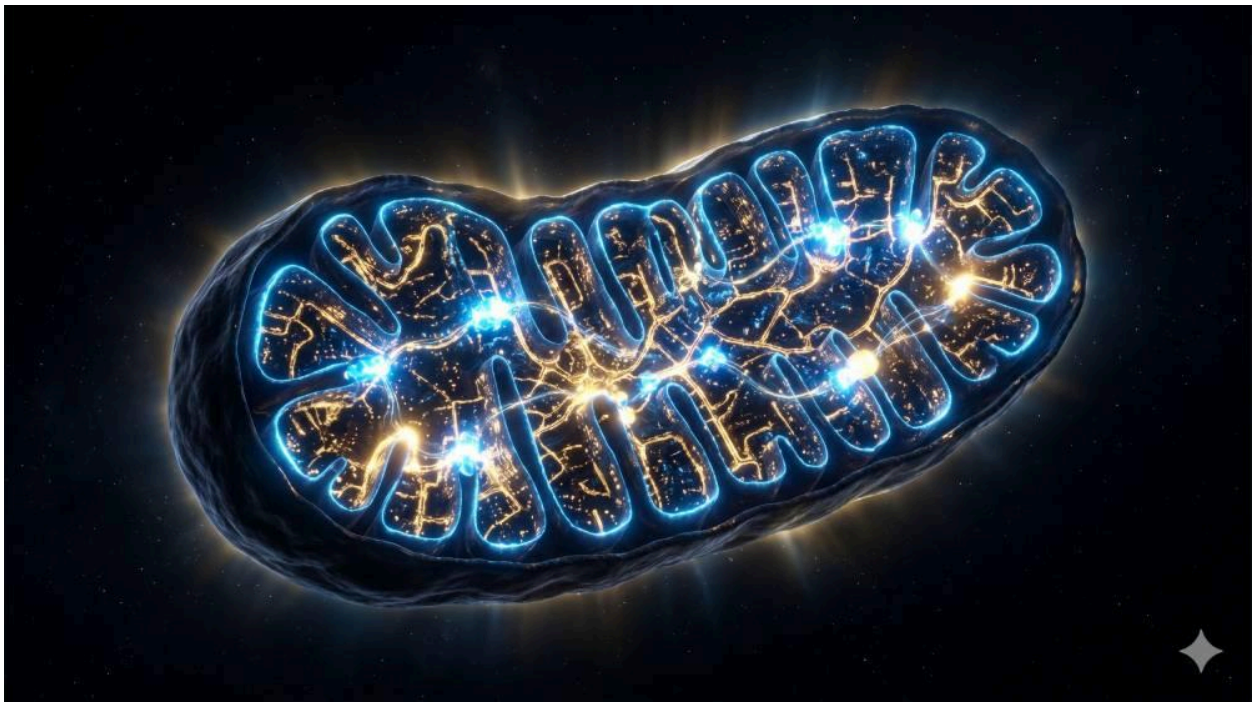
The Anesthesia Experiment and Its Implications

Lane's early experimental work makes this more than speculation. In infrared spectroscopy studies of brain tissue under anesthesia, he observed something unexpected: when anesthetic concentrations in the blood fluctuate, the redox state of cytochrome oxidase — that is, the balance of electron flow through the terminal enzyme of the respiratory

chain, deep in the mitochondria — shifts measurably. The respiratory chain is disrupted. Electrons that should flow through it and be handed cleanly to oxygen are instead dissipated as free radicals.

This is not what happens when you simply inhibit a neurotransmitter receptor. This is a disruption of the fundamental energy machinery of the cell — occurring in the mitochondria, preceding and accompanying the loss of consciousness.

The significance is layered. First, it suggests that what anesthetics primarily do is not silence neurons but destabilise the electrochemical gradient that is the physical correlate of feeling. Second, it explains why such chemically diverse compounds can all produce anesthesia: they share not a structural similarity but a functional one — they all interfere, through various mechanisms, with electron flow and proton gradients in the mitochondrial inner membrane. Third, and most provocatively, it implies that the amoeba rendered motionless by chloroform has lost something analogous to what we lose under general anesthesia — not neural computation, since it has none, but the electrochemical state that is the substrate of its most primitive awareness of its own condition.



Where the Line Falls — and Whether It Should

Lane frames his hypothesis around motile organisms, and with good reason. For natural selection to act on a feeling, that feeling must make a difference to what the organism does. The most obvious way a primitive "not OK" signal makes a difference is by driving movement — withdrawal from a noxious stimulus, approach toward a nutrient source. Motility gives feeling a clear functional outlet.

But the line between motile and fixed organisms is less sharp than it first appears. Consider what growth actually is: a root tip growing toward water is not random — it is directional, responsive, and driven by continuous sampling of the chemical environment. Differential cell elongation on one side of the root tip produces curvature toward the stimulus. The mechanism is utterly different from muscular locomotion, but the functional logic is identical: an internal representation of condition driving directed movement through space. The movement is simply slow, and it builds new tissue rather than displacing existing tissue. That is a difference of mechanism, not of principle.

Reproduction, similarly, is motility in time rather than space. An organism projects itself forward, and that projection is modulated by its assessment of its own condition. A bacterium that sporulates under stress is acting on its "not OK" signal as surely as one that swims away from a toxin — it is extending its temporal reach, doing something with its internal state that natural selection can act on. Metabolism itself, the most universal property of life, is directed activity at the molecular scale — perpetually responsive to internal condition, perpetually adjusting fluxes of energy and matter in ways driven by the cell's electrochemical state.

If the criterion for feeling is that an internal electrochemical state makes a difference to what an organism does, then growth, reproduction, and metabolism all qualify. And all three are universal properties of life.

Which leads to a conclusion that Lane's framework implies, even if he has not stated it in these terms: **feeling may be as universal a property of life as metabolism itself — because it is metabolism, apprehended from the inside.** Feeling may be metabolism itself — the same electrochemical process the biochemist measures with instruments, but experienced from within rather than observed fromwithout.

This is not mysticism. It requires no violation of parsimony, no appeal to forces beyond the known physics of electrochemistry. It requires only taking seriously the proposal that the proton-motive force, which is universal in cellular life, is the physical substrate of the most primitive feeling — and following that proposal to its logical conclusion.

A Note on Parsimony

Any serious treatment of the origin of consciousness must reckon with the principle of parsimony — Occam's razor — the demand that we not multiply explanatory entities beyond what the evidence requires. This principle is not optional in science. It is constitutive of it, in the same way that the rules of chess are not optional features of the game but the conditions of its possibility.

The great temptation in consciousness studies is to reach for rich cognitive vocabulary — problem-solving, goals, preferences, memory, mind — and apply it liberally to systems that exhibit adaptive behaviour. This approach has a certain narrative appeal. It is also, without rigorous constraint, scientifically empty. If every adaptive response is evidence of cognition, then cognition explains nothing, because it excludes nothing.

The discipline Lane exercises — asking what the physical substrate is, what its evolutionary function is, what you would need to measure to test the hypothesis — is precisely what keeps his speculation from becoming mere storytelling. The proton-motive force is not a metaphor. It is a specific, measurable, manipulable physical quantity. The

prediction that disrupting it should correlate with the loss of behavioural signs of awareness — across organisms, independently of whether they have neurons — is testable. It has already received preliminary support from the anesthesia experiments. That is how science is supposed to work.

Lane's framework cannot bridge the final gap — the hard problem, the question of why any physical state should be accompanied by subjective experience rather than simply occurring in the dark. No current theory can. But he has done something valuable and rare: he has identified the most ancient, most universal, most physically specific candidate for the substrate of feeling, and shown that it is in principle amenable to empirical investigation. He has transformed a purely philosophical puzzle into a partly scientific research program while remaining honest about where the science ends and the mystery begins.

That is playing by the rules. And in science, playing by the rules is not a limitation. It is the only way the game produces knowledge.

Conclusion: Two Billion Years of Feeling

The origin of consciousness, on the account developed here, is not a moment in the evolution of animal brains. It is a moment — if we can call it that — in the evolution of the cell: the acquisition, roughly two billion years ago, of mitochondria, and with them the steep electrochemical gradients that are the physical signature of metabolic health or distress.

Every living cell that respires maintains this gradient. Every disruption of that gradient — by starvation, by poison, by anesthetic — is a disruption of the most primitive feeling. The elaboration of that feeling into the rich interiority of animal consciousness, into emotion and pain and pleasure and self-awareness, required nervous systems, brains, and hundreds of millions of years of further evolution. But the root — the

bare physical fact of an electrochemical state that registers the cell's own condition — is as old as complex cellular life itself.

Consciousness, then, is not what the brain does. It is what life does, whenever it is organised well enough to feel the difference between doing well and doing badly. The brain does it with extraordinary sophistication. But it did not invent it. It inherited it, from bacteria, two billion years ago, along with the mitochondria that still power every thought we think.

That inheritance is worth taking seriously. Not because it dissolves the mystery of consciousness — it does not — but because it relocates the mystery where it belongs: not in the complexity of neural circuits, but in the ancient, universal, still poorly understood chemistry of life itself.

By Nick Lane:

- **Transformer: The Deep Chemistry of Life and Death** (2022)
- **The Vital Question: Why Is Life the Way It Is?** (2015)
- **Life Ascending: The Ten Great Inventions of Evolution** (2009)
- **Power, Sex, Suicide: Mitochondria and the Meaning of Life** (2005)

Lynn Margulis (Sagan) **"On the origin of mitosing cells."** *Journal of Theoretical Biology*, 14(3), pp. 225–274 (1967).